

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPerch_un n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Sun Mar 7 02:33:19 1999; MasPar time 389.26 Seconds
1390.239 Million cell updates/sec
Tabular output not generated.

Title: >US-08-962-094-1
Description: (1-229) from US08962094.seq
Perfect Score: 223
N.A. Sequence: 1 CGGCTCGAGCTCTTAGGCTT.....TCGNACCACTGGGACCACTG 229
Comp: GCCGAGCTCGAGATCCGAA.....ACGNTGTGACGCTGTGTGAC

Scoring table: TABLE default
Gap 6

*Nmatch STD : Dbase 0; Query 0

Searched: 602357 segs, 1181590623 bases x 2

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: emb157
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_vl
Database: genbank110
16:gb_bal 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl 25:gb_pi2 26:gb_pi1
27:gb_pi2 28:gb_pi3 29:gb_ro 30:gb_sts 31:gb_sts 32:gb_sy
33:gb_un 34:gb_vl

Statistics: Mean 9.529; Variance 5.152; scale 1.850

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

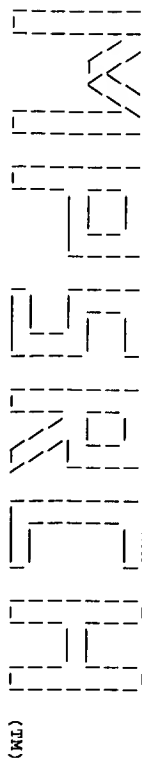
SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description	Pred. No.
No.								

No matches found.

Search completed: Sun Mar 7 02:44:29 1999
Job time : 670 secs.

Arthur L.
08/06/2004 Page 1
Seq-ID# 17-20
1-5 & 17-20



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_mn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 02:54:43 1999; Maspar time 43.34 Seconds

718.825 Million cell updates/sec
Tabular output not generated.

Title: >US-08-962-094-1
Description: (1-229) from US08962094.seq
Perfect Score: 223
N.A. Sequence: 1 CGGCTCGAGCTCTAGGCTT.....TGCNACGACTGCGACACTG 229
Comp: GCCGAGCTCGAATCCGA.....ACGNTGCTGAAGCTGTGAC

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0: Query 0

Searched: 188442 segs, 68026449 bases x 2

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 7.735; Variance 5.150; scale 1.502

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.

No matches found.

Search completed: Sun Mar 7 02:58:23 1999
Job time : 220 secs.

		RESULT	4		
D _b	122	GCTGACACTATTCAGCGAATCMGGTCCMGCGATGAAGAACCCTCATGTCTAAACCACT	181		
Oy	145	GCTGCACACGNATTCAGCAGTACTGTGGTTCTTGATGATGAAGGCCCTGANCCTGAACAACACT	204		
D _b	182	GCTGCTGCACACCACTGCAGCACACTg	206		
Oy	205	GCTGCTGCACNACCACCTGCAGCACACTg	229		
LOCUS		w72838	432 bp	mRNA	EST
DEFINITION		d257g12.s1 Soares fetal heart NDHH19W Homo sapiens cDNA clone 34406_3' similar to contains element MER40 repetitive element ; mRNA sequence.			16-OCT-1996
ACCESSION	NID	w72838			
KEYWORDS	SOURCE	g1382814			
ORGANISM		human .			
REFERENCE	AUTHORS	Homo sapiens			
		Eukaryota; Eukarya; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 432)			
		Hillier,L., Clark,N., Dubucque,T., Elliston,K., Hawkins,M., Holmen,W., Hultman,M., Kuecaba,T., Le,M., Lennon,G., Maira,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.			
TITLE	JOURNAL	The Washo-Merck EST Project			
COMMENT		Unpublished (1995)			

Contact: Wilson R

```

WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 508 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 394.
Location/Qualifiers
1. 432
/organism="Homo sapiens"
/note="Organ: heart; Vector: pUT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTTACCATCTGAGATGGAGGCGGCCGCATCTTTTATTTTATTTT 3'], R
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pUT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHL19W."
/db_xref="taxon:9606"
/clone_3="344806"
/clone_11b="Soares fetal heart NBHL19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>432)
BASE COUNT      128 a      78 c      124 g      102 t
ORIGIN
Query Match      75.8%; Score 169; DB 21; Length 432;
Best Local Similarity 96.7%; Pred.No. 4,77e-260;
Matches 175; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
DB 252 CAGTGTGTCGACGATGTTTCACACACAGATGTTTATGAGATCAGAGGGCGCTTCATCATCAGACAG 311

```

Cp	229	CAGTGGTGGCAGTGGTGCACACACACATGGTTTCACACNCCAGGGCTTCATCATCAGACAG	170
Db	312	GACCACTAAGCTGGATACGTGCAGTGCAGTGGAGCAGCTGTTGCGGATTTCTGGGCGAGAGCC	371
Cp	169	GACCAGTA -GCTGGATNGTGTCTAGCTGNNAGCAGCTGTTGGATTTCTGGGCGAGAGCC	111
Db	372	AGAAAGATGGAAACCTCCCAAGAGTACCGACGACGCTAAAGACTTATGGTGGTGACATGA	431
Cp	110	AGAAAGATGGAAACCTCCCAAGAGTACCGACGACGCTAAAGACTTATGGTGGTGACATGA	51
Db	432	A 432	
Cp	50	A 50	

RESULT	5	443 bp	mRNA	EST	11-SEP-1998
LOCUS	AI177172				
DEFINITION	gb99aa08.x1 Scores fetal heart NDHH19W Homo sapiens cDNA clone IMAGE:1708214 3', mRNA sequence.				
ACCESSION	AI177172				
NID	93595686				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
REFERENCE	1 (bases 1 to 443)				
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP)				
JOURNAL	Tumor Gene Index				
COMMENT	Unpublished (1997)				

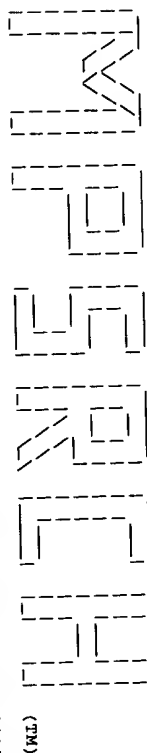
~~Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m13 fwd, RT from Amer sham.
Location/Qualifiers
1. .443 chr2:100000000~~

```

/organism="mouse sapiens" Vector: pT73D (Pharmacia) with a
/notest="Organ: heart; Site1: Not I; Site2: Eco RI; 1st
modified polylinker; Site3: Not I; Site4: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [
TGGTACATCTGAGAGTGAGAGCGCCGATCTTTTCTTTTCTTTTCTTTT
3', 1,
double-stranded cDNA was size selected, ligated to Eco R
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBH19W."
/db_xref="taxon:9606"
/clone_image="IMAGE:1708214"
/clone_id="Soares_fetal_heart_NBH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
129 a 80 c 129 g 103 t 2 others

```

Query Match	71.3%	Score 159;	DB 17;	Length 443;
Best Local Similarity	95.3%;	Pred. No. 3,48e-241;		
Matches 162;	Conservative	0;	Mismatches 8;	Indels 0;
				Gaps 0;
<hr/>				
Db	274	CAGTGCACAGTGGTTCCACAGCAGTGGTTTCAGACATCAGGGGCTTCATCATCAGCAG	333	
Cp	229	CAGTGGTCCAGTGCAGTGCAGAGAGAGTGGTTTCAGACATCAGGGGCTTCATCATCAGCAG	170	
Db	334	GACCACTAGCTGATACCTGTGTACGCTGGACAGCAGTGTTCGATTCCTGGCAGAGACCA	393	



 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mparc_hun n.a. - n.a. database search, using Smith-Waterman algorithm
 Run on: Sun Mar 7 02:58:46 1999; Maspar time 514.57 Seconds
 1414.510 Million cell updates/sec
 Tabular output not generated.

Title: >US-08-962-094-2
 Description: (1-308) from US08962094.seq
 Perfect Score: 308
 N.A. Sequence: 1 TAGGCTTTGAAGCATTTTG.....TTGGGATCTTCGATAGT 308
 Comp: ATCCGAACCTTCGTAAAC.....AACCCTAGAAGGCTTACCA

Scoring table: TABLE default
 Gap 6

Nmatch STD: Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 50%
 Listing first 1000 summaries

Database: emb157
 Item: ba 2:em_fun 3:em_htg 4:em_hun1 5:em_hun2 6:em_in
 7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
 13:em_ro 14:em_un 15:em_v1
 Database: genbank110
 16:gb_ba1 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_ov
 22:gb_pat 23:gb_ph 24:gb_pl 25:gb_pl2 26:gb_pl1
 27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
 33:gb_un 34:gb_v1

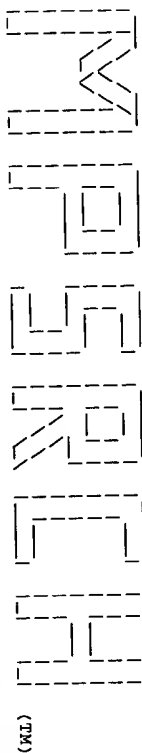
Statistics: Mean 9.810; Variance 5.472; scale 1.793
 Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred. No.

No matches found.
 Search completed: Sun Mar 7 03:11:01 1999
 Job time : 735 secs.

12



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpserch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 03:23:57 1999; Mspar time 58.14 Seconds
720.802 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-2
Description: (1-308) from US08962094.seq
Perfect Score: 308 1 TAGGCTTGACGATTTTG.....TTGGGATCTTCGATGT 308
N.A. Sequence: ATCCGAACTTCGTAAAC.....AACCCTAGAGCTTACCA
Comp:

Scoring table: TABLE default
Gap 6

* Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 508
Listing first 1000 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 7.972; Variance 5.293; scale 1.506

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description	Pred. No.

~~No matches found.~~

Search completed: Sun Mar 7 03:27:25 1999
Job time : 208 secs.

 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPSEARCH n.a. - n.a. database search, using Smith-Waterman algorithm
 Run on: Sun Mar 7 03:11:25 1999; Maspar time 436.14 Seconds
 1264.640 Million cell updates/sec

Tabular output not generated.
 Title: >US-08-962-094-2
 Description: (1-308) From US08962094.seq
 Perfect Score: 308
 N.A. Sequence: 1 TAGGCTTGAAGCATTGTTG.....TTGGGAGCTTCGAGATGGT 308
 Comp: ATCCGAACTTCGTAAAC.....AACCCCTGAAGAGCTTACCA

Scoring table: TABLE default
 Gap 6

Mismatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 50%
 Listing first 1000 summaries

Database: emb1-est156
 1:em-est1-2:em-gss1 3:em-gss2 4:em-gss3
 genbank-est1109
 5:gb-est1 6:gb-est10 7:gb-est11 8:gb-est12 9:gb-est13
 10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17
 14:gb-est18 15:gb-est19 16:gb-est20 17:gb-est21
 18:gb-est22 19:gb-est23 20:gb-est24 21:gb-est25 22:gb-est26
 23:gb-est27 24:gb-est28 25:gb-est29 26:gb-est30 27:gb-est31
 28:gb-gss3 29:gb-gss4

Statistics: Mean 9.930; Variance 2.132; scale 4.658

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	293	95.1	403	16	R75793	Y121f09.r1 Homo sapien
2	293	95.1	436	21	W72837	z457g12.r1 Soares feta
3	291	94.5	294	25	AA340069	EST45219 Fetal skin Ho
4	262	85.1	432	21	W72838	z457g12.s1 Soares feta
5	251	81.5	443	17	A1127172	q099608.x1 Soares feta
6	232	75.3	402	18	A1139456	q020901.x1 Soares feta

Note: Post-processor removed 994 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1

LOCUS R75793 403 bp mRNA EST
 DEFINITION Y121f09.r1 Homo sapiens cDNA clone 158921 5' similar to contains
 LTR3 repetitive element ;
 R75793
 ACCESSION 9850475
 NID
 KEYWORDS
 SOURCE

ORGANISM

REFERENCE
 AUTHORS

TITLE
 JOURNAL
 COMMENT
 The WashU-Merck EST Project
 Unpublished (1995)

Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 High quality sequence strops: 307
 Source: IMAGE Consortium, LIND
 This clone is available royalty-free through LIND; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

source

BASE COUNT

82 a 113 c 89 g 116 t 3 others

ORIGIN

Query Match 95.1% Score 293; DB 16; Length 403;
 Best Local Similarity 99.3% Pred. No. 0.00e+00;
 Matches 294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db	2	GCATTTTGTCTGTCGCCCTGATCTTCAGTCCACACCATGAAATCTTACGATCCG 61
Yy	12	GCATTTTGTCTGTCGCCCTGATCTTCAGTCCACACCATGAAATCTTACGATCCG 71
Db	62	GNACTCTGGGAGTTCCATCTTCTGCTCCGACGATCCGACGACGATCTGCTCA 121
Yy	72	GTAATCTTGGGAGTTCCATCTTCTGCTCCGACGATCCGACGATCTGCTCA 131
Db	122	GGGACAGTATCCAGTCTGCTCTGATGATGAAGCCCTGATGCTGAACCACT 181
Yy	132	GCTGACAGTATCCAGTCTGCTCTGATGATGAAGCCCTGATGCTGAACCACT 191
Db	182	GCTGTCGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 241
Yy	192	GCTGTCGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 251
Db	242	ACTGCTGTAAGACATTCAGTTTACCAATGGGTTGGGATCTCCCAATGG 297
Yy	252	ACTGCTGTAAGACATTCAGTTTACCAATGGGTTGGGATCTCCCAATGG 307

	RESULT	2	
	LOCUS		
	DEFINITION	W72837	436 bp mRNA
		Z657912.r1 Soares fetal heart NbH119g Homo sapiens cDNA clone	EST 16-OCT-1996
		344806.5' similar to contains element MBR40 repetitive element ;	
	ACCESSION	W72837	
	NID	91382813	
	KEYWORDS		
	SOURCE	EST.	
	ORGANISM	human.	
		Homo sapiens	
REFERENCE		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;	
AUTHORS		Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
		1 (bases 1 to 436)	
		Hillier,L., Clark,N., Dubque,T., Elliston,K., Hawkins,M.,	
		Holtman,M., Holtman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,	
		Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,	
		Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and	
		Wilson.R.	
JOURNAL COMMENT		The Mashu-Merck EST Project	
		Unpublished (1995)	
TITLE		Contact: Wilson RK	
JOURNAL		Mashu-Merck EST Project	
COMMENT		Washington University School of Medicine	
		4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
		Tel.: 314 286 1800	
		Fax: 314 286 1810	
		Email: est@watson.wustl.edu	
		This clone is available royalty-free through LNL ; contact the	
		IMAGE Consortium (info@image.llnl.gov) for further information.	
		Insert length: 508 Std Error: 0.00	
		Seq primer: mod.REGA+ET.	
FEATURES		Location/Qualifiers	
source		1..436	

Query Match	95.1%	Score 293	DB 21	length 436
Best Local Similarity	99.3%	Pred. No. 0.00e+00		
Matches	295	Conservative	0	Mismatches 2; Indels 0; Gaps 0
Db	1	GCATTTTGTGCTGCTGCTCCGATCTTATGTCACCAACATGAAGTTCTTACAGCTCGT	60	
Qy	12	GCATTTTGTGCTGCTGCTCCGATCTTACGTCACCAACATGAAGTTCTTACAGCTCGT	71	
Db	61	GTACTTTGGGAGTTTCCATCTTTCTGTGCTCTGTCGCCAGATTCGACACAGCTGTCCA	120	
Qy	72	GTACTTTGGGAGTTTCCATCTTTCTGTGCTCTGTCGCCAGATTCGACACAGCTGTCCA	131	
Db	121	GCTGACACGTATCCAGCTACTGTCGTCGTCGATGATGAAGCCCTGATGTGAACCACT	180	
Qy	132	GCTGACACGTATCCAGCTACTGTCGTCGTCGATGATGAAGCCCTGATGTGAACCACT	191	

D_b 181 GCTGCTGCACAGCAGTGTGGAGCCACTGTCCTCTCTTACATGCATCGAACCAACGCCGTGCTTTACC 240
Q_y 192 GCCTCTGCACACAGCTGCGAGCCAGCTGCTCTCTCTCTCTCACATCGAACCAACGCCGTGCTTTACC 251
D_b 241 ACTGCTGTTAAAGACATTCCAGTTTTCACCAAAATGGGTTGGGGAATCTCCGCAATGGT 297
Q_y 252 ACTGCTGTTAAAGACATTCCAGTTTTCACCAAAATGGGTTGGGGAATCTCCGCAATGGT 308

RESULT	3	294 bp	MRNA	EST	21-APR-1997
LOCUS	AA340069				
DEFINITION	EST45219 Fetal skin Homo sapiens CDNA 5' end similar to				
ACCESSION	AA340069				
NID	g1992307				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Eukaryotes; Metazoa; Chordata;				

REFERENCE	AUTHORS
1 (bases 1 to 294)	
	Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Feldner,R.A.,
	Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
	White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Well,C.,
	Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
	Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.V., Geoghegan,N.S.,
	Gloder,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
	Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
	Moreno-Palmanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
	Phillips,C.A., Ryder,S.E., Scott,J.L., Seudek,D.M., Shirley,R.,
	Smell,L.K.V., Spriggs,T.A., Utechtack,T.R., Weidman,J.F., Li,X.,
	Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T., O'Malley,J.F.,
	Dimke,D., Feng,D.F., Ferris,A., Fischer,C., Hastings,G.A.,
	He,W.M., Hu,J.S., Greene,J.M., Gruber,C., Hudson,P., Kim,A.K.,
	Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
	Royak,L., Wei,Y.F., Wingard,J., Xu,C., Yu,G.L., Ruben,S.M.,
	Dallion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
	Fraser,C.M. and Venter,J.C.
TITLE	Initial assessment of human gene diversity and expression patterns
JOURNAL	based upon 83 million nucleotides of cdna sequence
MEDLINE	Nature 377 (6547 Suppl.), 3-174 (1995)
COMMENT	96026280
OTHER ESTS	TC0117518

The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavett@igf.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/cdb/hgi.html>).
Seq primer: M13 Reverse.

FEATURES	Location/Qualifiers
SOURCE	1. 294
	/organism="Homo sapiens"
	/note="Organ: skin; Vector: pBluescript SK-; Site:1;
	ECORI; Site:2: XhoI; Vector: pBluescript SK-; Site:1;
	/db_xref="ATCC (inhost):141700"
	/db_xref="taxon:9606"
	/clone_lib="fetal skin"
	/tissue_type="epithelium"
	/cell_type="epithelial cell"
	/dev_stage="fetus"
MRNA	<1. >294
BASE COUNT	60 a 91 c 58 g 82 t 3 others
ORIGIN	
Query Match	94.5%; Score 291; DB 25; Length 294;
Best Local Similarity	99.0%; Pred. No. 0.00e+00;
Matches 291; Conservative	0; Mismatches 3; Indels 0; Gaps 0

Db 1 CTTTGAAGCATTTTNTCTGTCCTGATCTTCAGGTACACACCAATGAAGTCTTAGC 60
 CTTTGAAGCATTTTNTCTGTCCTGATCTTCAGGTACACACCAATGAAGTCTTAGC 64
 5 CTTTGAAGCATTTTNTCTGTCCTGATCTTCAGGTACACACCAATGAAGTCTTAGC 64
 Oy 61 AGTCTGTGACCTCTGGAGGTTTCATCTTTGCTCTGGCCAGAAATCCAGAACAGC 120
 65 AGTCTGTGACCTCTGGAGGTTTCATCTTTGCTCTGGCCAGAAATCCAGAACAGC 124
 Db 121 TGTCTCAGCTGACAGCTATCCAGCTACTGCTCTGATGATGAAGCCCTGATGCTGA 180
 125 TGTCTCAGCTGACAGCTATCCAGCTACTGCTCTGATGATGAAGCCCTGATGCTGA 184
 Oy 181 AACCACTGCTGTGTCGAACCACTGCGACCACTGCTGCTCTACCACTGCAACCCGCTGC 240
 185 AACCACTGCTGTGTCGAACCACTGCGACCACTGCTGCTCTACCACTGCAACCCGCTGC 244
 Db 241 TTCTTACCACTGCTGTGTCGAACCACTGCTGCTCTACCACTGCAACCCGCTGC 294
 245 TTCTTACCACTGCTGTGTCGAACCACTGCTGCTCTACCACTGCAACCCGCTGC 298

RESULT 4 W72838 432 bp mRNA EST 16-OCT-1996
 LOCUS 2d57912.s1 Soares fetal heart NBH19W Homo sapiens cDNA clone
 DEFINITION 344806 3' similar to contains element MER40 repetitive element ;,
 mRNA sequence.
 W72838
 ACCESSION 91382814
 NID EST.
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Eukaryotes; Eukaryotes; Metazoa; Chordata;
 Vertebrata; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 432)
 REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, R., Hawkins, M.,
 Holman, M., Hultman, M., Kuchba, T., Le, M., Lennon, G., Mearns, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevaaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
 Wilson, R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 508 Std Error: 0.00
 Seq primer: mob.REGA+ET
 High quality sequence stop: 394.
 Location/Qualifiers
 1.432
 /organism="Homo sapiens"
 /note="Organ: heart; Vector: p773D (Pharmacia) with a
 modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAGAGGAGGAGGCGGCGCATCTTTTCTTTTCTTTT 3']
 TGTACCAATCTGAGAGGAGGAGGCGGCGCATCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p773 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by
 M.Fatima Bonaldo. This library was constructed from the
 same fetus as the fetal lung library, Soares fetal lung
 NBH19W."
 /db_xref="taxon:9606"
 /clone="344806"
 /clone_1lb="Soares fetal heart NBH19W"

/sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 complement(<1..>432)
 BASE COUNT 128 a 78 c 124 g 102 t
 ORIGIN
 Query Match 85.1%; Score 262; DB 21; Length 432;
 Best Local Similarity 98.9%; Pred. No. 0.00e+00;
 Matches 270; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 Db 160 ACCATGCGGAGATCCCAACCCATTGGGTAAACGTGATCTTTAGACAGCTGCT 219
 Cp 308 ACCATGCGGAGATCCCAACCCATTGGGTAAACGTGATCTTTAGACAGCTGCT 249
 Db 220 AGAAGCAGCGGTGTTGTCAGTGTAGAGCAGCAGTGTGCGAGTGTGACAGCAGT 279
 Cp 248 AGAAGCAGCGGTGTTGTCAGTGTAGAGCAGCAGTGTGCGAGTGTGACAGCAGT 189
 Db 280 GGTTCAGCATCAGGGCTTCATCATCAGCAGCAGTAACTGATAGCTGCTGAGCTG 339
 Cp 188 GGTTCAGCATCAGGGCTTCATCATCAGCAGCAGTAACTGATAGCTGCTGAGCTG 130
 Db 340 GAGCAGCTGTGTCGATTCGAGTGTGAGCAGCAGCAGTAACTGATAGCTGCTGAGCTG 399
 Cp 129 GAGCAGCTGTGTCGATTCGAGTGTGAGCAGCAGCAGTAACTGATAGCTGCTGAGCTG 70
 Db 400 GGAAGCTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
 Cp 69 GGAAGCTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 37

RESULT 5 A1127172 443 bp mRNA EST 11-SEP-1998
 LOCUS gp99a08.x1 Soares fetal heart NBH19W Homo sapiens cDNA clone
 DEFINITION IMAGE:1708214 3', mRNA sequence.
 A1127172
 ACCESSION 93595686
 NID EST.
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Euthera;
 Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 443)
 REFERENCE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS Tumor Gene Index
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40ml3 fwd. ET from Amersham.
 Location/Qualifiers
 1.443
 /organism="Homo sapiens"
 /note="Organ: heart; Vector: p773D (Pharmacia) with a
 modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAGAGGAGGAGGCGGCGCATCTTTTCTTTTCTTTT 3']
 TGTACCAATCTGAGAGGAGGAGGCGGCGCATCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p773 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by
 M.Fatima Bonaldo. This library was constructed from the
 same fetus as the fetal lung library, Soares fetal lung
 NBH19W."
 /db_xref="taxon:9606"
 /clone="IMAGE:1708214"
 /clone_1lb="Soares fetal heart NBH19W"

Mon Mar 8 15:41:59 1999

US-08-962-094-3.rge

MPERCH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPERCH_un n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 03:27:49 1999; Maspar time 347.59 Seconds
1339.364 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-3
Description: (1-197) from US08962094.seq
Perfect Score: 197
N.A. Sequence: 1 GTTTACCCCAATGGCTGG.....AAATACTATGACACAT 197
Comp: CAAATGGGTTTACCAAC.....TTTATGTATGACTGCTGTA

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 602357 segs, 1181590623 bases x 2

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: emb157
1:em_ba 2:em_fun 3:em_htg 4:em_hun1 5:em_hun2 6:em_in
7:em_om 8:em_ov 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_yi
Database: gendank110
16:gb_ba1 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl1 25:gb_pl2 26:gb_pl3
27:gb_pr1 28:gb_pr2 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_yi

Statistics: Mean 9.496; Variance 5.472; scale 1.735

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

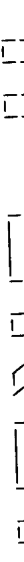
Result No.	Query No.	Score	Match	Length	DB	ID	Description	Pred. No.
------------	-----------	-------	-------	--------	----	----	-------------	-----------

No matches found.
Search completed: Sun Mar 7 03:38:19 1999
Job time : 630 secs.

ArThuy, L
08/06/2004
Seq. IDs 1-20
Page 1

24

25



```
MPsrch_nn      n.a. - n.a. database search, using Smith-Waterman algorithm
```

```
Run on:      Sun Mar 7 03:48:19 1999;  MasPar time 38.11 seconds
              703.291 Million cell updates/sec
```

Tabular output not generated.

```

Title: >US-08-962-094-3
Description: (1-197) from US08962094.seq
Perfect score: 197
N A. Sequence: 1 GTTTTACCCAAATGGGTGG.....AAAATACATATGACACAT 197
Comp: CAAATGGGTTTACCAAC.....TTTATTGATCTCTGTGTA

```

Scoring table: TABLE default
Gap 6

```
Nmatch      STD :      Dbase 0;      Query 0
```

Searched: 188442 seqs, 68026449 bases x 2

post-processing:	Minimum Match 50%
	Listing first 1000 summaries

```
Database:
n-genseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40
```

Statistics: Mean 7.570; Variance 4.834; scale 1.566

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match Score	Length	DB	ID	Description	Pred. No.

No matches found.

Search completed: Sun Mar 7 03:51:06 1999
Job time : 167 secs.

2

3

KEYWORDS

EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

EST.
human.
Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 432)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 508 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 394.
Location/Qualifiers
1. 432

FEATURES

SOURCE

/organism="Homo sapiens"
/note="Organ: heart; Vector: p773D (Pharmacia) with a
modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAGTGGAGCGCGCCGATCTTTTCTTTTCTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M. Fatima Boudado. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHL19W."
/db_xref="taxon:9606"
/clone_lib="Soares fetal heart NBHL19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 128 a 78 c 124 g 102 t
ORIGIN

Query Match 99.0%; Score 195; DB 21; Length 432;
Best Local Similarity 100.0%; Pred. No. 1,71e-267;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GTTGCTCATAGTATTTTATTTGAAAGAAATAAAGTATGAGATTAAGGGATAT 60
|||||
Cp 195 GTTGCTCATAGTATTTTATTTGAAAGAAATAAAGTATGAGATTAAGGGATAT 136
|||||
Db 61 CGTAGCAGAGTAACTAGTGGATGAATACAGAGACATGAATAGTGTGACAAATG 120
|||||
Cp 135 CGTAGCAGAGTAACTAGTGGATGAATACAGAGACATGAATAGTGTGACAAATG 76
|||||
Db 121 CAGAGACTACAGTATCATCTCAGGACACACTTACATTTGGGAGATCCCAAC 180
|||||
Cp 75 CAGAGACTACAGTATCATCTCAGGACACACTTACATTTGGGAGATCCCAAC 16
|||||
Db 181 CCATTGGGTAAC 195
|||||
Cp 15 CCATTGGGTAAC 1

RESULT 3

LOCUS W72837 436 bp mRNA EST 16-OCT-1996

DEFINITION zds7g12.1 Soares fetal heart NBHL19W Homo sapiens cDNA clone
344806 5' similar to contigs element-MR40-repetitive element ;
RNA sequence.

ACCESSION W72837
NID 9182813
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 436)
Vertebrata; Euthera; Primates; Catarrhini; Homnidae; Homo.
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 508 Std Error: 0.00
Seq primer: mob.REGA+ET.
Location/Qualifiers
1. 436

FEATURES

SOURCE

/organism="Homo sapiens"
/note="Organ: heart; Vector: p773D (Pharmacia) with a
modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAGTGGAGCGCGCCGATCTTTTCTTTTCTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M. Fatima Boudado. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHL19W."
/db_xref="taxon:9606"
/clone_lib="Soares fetal heart NBHL19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 92 a 128 c 81 g 134 t 1 others
ORIGIN

Query Match 85.3%; Score 168; DB 21; Length 436;
Best Local Similarity 99.4%; Pred. No. 4.86e-223;
Matches 174; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 262 GTTTTACCAATGGGTTGGGATCCGAAATGAGAGTGTGCTCGATGGAATC 321
|||||
Cp 1 GTTTTACCAATGGGTTGGGATCCGAAATGAGAGTGTGCTCGATGGAATC 60
|||||
Db 322 ACCTTGAGTCTTTCGCAATNGTCAACACTATTCATGCTTCGATTCATCACT 381
|||||
Cp 61 AGCTTGAGTCTTTCGCAATNGTCAACACTATTCATGCTTCGATTCATCACT 119
|||||
Db 382 ACTTACCTGCTTACGATATCCCTTATCTCTATTCAGTTATTTCTTCAAA 436
|||||
Cp 120 ACTTACCTGCTTACGATATCCCTTATCTCTATTCAGTTATTTCTTCAAA 174

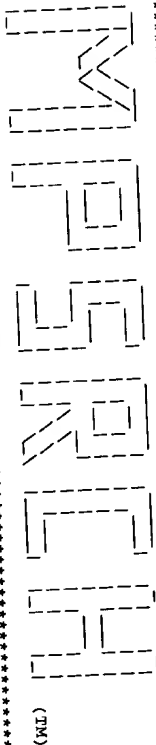
Mon Mar 8 15:42:01 1999

Search completed: Sun Mar 7 03:47:54 1999
Job time : 552 secs.

US-08-962-094-3.rst

Mon Mar 8 15:42:02 1999

US-08-962-094-4.rge



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_un n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Sun Mar 7 03:51:30 1999; Maspar time 778.00 Seconds
1464.084 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-4
Description: (1-482) from US08962094.seq
Perfect Score: 482
N.A. Sequence: 1 CGGCTCGAGCTCTTAGCCTT.....AAATTAATGATGACACAT 482
Comp: GCCGAGCTCGAATCGAA.....TTTATGATACTGTTGTA

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

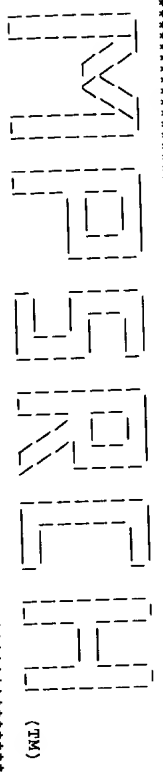
Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: emb157
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_com 8:em_ov 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_vl
Database: genbank110
16:gb_bal 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_com 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl1 25:gb_pl2 26:gb_pl3
27:gb_pr1 28:gb_pr2 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_vl

Statistics: Mean 10.452; Variance 5.845; scale 1.788
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description	Pred. No.
No matches found.					
Search completed: Sun Mar 7 04:09:11 1999					
Job time : 1061 secs.					



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_n n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Sun Mar 7 04:24:51 1999; MasPar time 86.95 Seconds
754.228 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-4
Description: (1-482) from US08962094.seq
Perfect Score: 482
N.A. Sequence: 1 CGGCTCGAGCTCTTGGCTT.....AAATACATATGAGCAACAT 482
Comp: GCCGAGCTCGAGATCCGAA.....TTTATTGATCTCGTTGTA

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: n.geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 8.488; Variance 5.519; scale 1.538

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

8			
Result	Query		
No.	Score	Match Length DB ID	Description
Pred. No.			

No matches found.

Search completed: Sun Mar 7 04:28:26 1999
Job time : 215 secs.

RESULT	2	EST	16-OCT-1996
LOCUS			
DEFINITION	2d57g12.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 344806.3, similar to contains element MERA0 repetitive element ; mRNA sequence.		
ACCESSION			
NID	W72838		
KEYWORDS	91382814		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Carnivhni; Hominiidae; Homo. 1 (bases 1 to 432)		
AUTHORS	Hillier, L., Clark, N., Dubuque, T., Elliston, R., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.		
TITLE	The WashU-Merck EST Project		
JOURNAL			
COMMENT	Unpublished (1995)		

```

/organism="Homo sapiens"
/Note="Organ: heart; Vector: p777D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - 0190(dt) primer [5
TCTTCCACATCTGACAGTGGGAGCGCCGACATCTTTTCTTTTCTT
3']"
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library was through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Donaldo. This library was constructed from the
same fecus as the fetal lung library, Soares fetal lung
NBH19W."
/db_xref="taxon:9606"
/clone="344806"
/clone_lib="Soares fetal heart NBH19W"
/sex="unknown"
/dev_stage="19 weeks"

```

Db	1	GTGGCCATATGTAATTTTATTTATTTGAAAAGAAATTAACCTATTAGAGATTAAGGGGATAT	6
Cp	480	GTTCCTCATATGTAATTTTATTTTATTTGAAAAGAAATTAACCTATTAGAGATTAAGGGGATAT	4
Db	61	CGTAGGCAAGGTAGTAGTGGATGGAATCAACAGAAAGCATGAATGTTGTACCAATTG	1
Cp	420	CGTAGGCAAGGTAGTAGTGGATGGAATCAACAGAAAGCATGAATGTTGTACCAATTG	3
Db	121	CATAAATCTCAAGCTGATTCATCTCAGGGACACACTACCTTACCTTGGGAGATCCCCAAC	1
Cp	360	CAGAAGACTCAAGCTGATTCATCTCAGGGACACACTACCTTACCTTGGGAGATCCCCAAC	3
Db	181	CCATTTGGGTAAATCGTAATGCTTTACAGACAGTGTAGAAAGCAGCGTGTTCAGT	2
Cp	300	CCATTTGGGTAAATCGTAATGCTTTACAGACAGTGTAGAAAGCAGCGTGTTCAGT	2
Db	241	GGTAGGACACAGCTGGTGCAGTGGTGGAGCAGCAGTGGTTTACATCAGAGGCTTC	3
Cp	240	GGTAGGACACACAGCTGGTGCAGTGGTGGAGCAGCAGTGGTTTACATCAGAGGCTTC	1
Db	301	ATCATCACACAGACCACTTAAGCTGGATACGTGTCAAGCTGGAGACAGCTTTGCGATTC	3
Cp	180	ATCATCACACAGACCACTTAAGCTGGATACGTGTCAAGCTGGAGACAGCTTTGCGATTC	1
Db	361	GGGAGAGACACAGAAAGATGAAATCCCAAGAGTACACAGACAGCTCTTAAGACTTCATGG	4
Cp	121	GGGAGAGAGACACAGAAAGATGAAATCCCAAGAGTACACAGACAGCTCTTAAGACTTCATGG	6
Db	421	TGGTGACATTAAT 432	
Cp	61	TGGTGACATTAAT 50	

FEATURES	source
LOCUS	3
DEFINITION	AI139456 402 bp mRNA
ACCESSION	q020901.x1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
NID	IMAGE:170192.3' similar to contains element MERR0 repetitive
KEYWORDS	AI139456
SOURCE	g3645428
ORGANISM	EST.
REFERENCE	human.
AUTHORS	Homo sapiens
TITLE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
JOURNAL	Primates; Catarrhini; Homiidae; Homo.
COMMENT	1 (bases 1 to 402)
	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgp .
	National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
	Tumor Gene Index
	Unpublished (1997)
	Contact: Robert Strausberg, Ph.D.
	Tel: (301) 496-1550
	Email: Robert.Strausberg@nih.gov
	This clone is available royalty-free through LNL; contact the
	IMAGE Consortium (info@image.llnl.gov) for further information.
	Seq primer: 40m13 fwd. ET from Amersham
	High quality sequence stop: 398.
	Location/Qualifiers
	1..402

```

FEATURES
    source          1..398
    Location/Qualifiers
        1..402
            /organism="Homo sapiens"
            /note="Organ: heart; Vector: p773D (Pharmacia) with a
            modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5

```

BASE COUNT
ORIGIN

KEYWORDS
SOURCE

ORGANISM

COMMENT

FEATURES

ORIGIN

Matches 3

QY

DEFINITION

REFERENCE

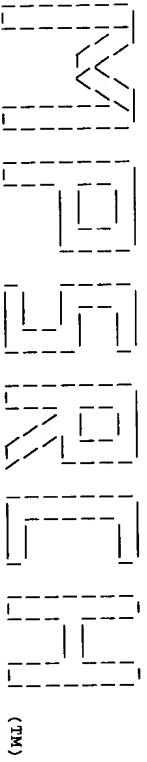
REFERENCE 1 (bases 1 to 294)

Mon Mar 8 15:42:04 1999

US-08-962-094-4.rst

Cp 143 GGAGCAGCTGTTGTCGGATTCTGGGAGAGAGACGAAAGATGGAACCTCCCAAGAGTACC 84
Db 420 AGGACTGCTAAGAACTTATGCTG 443
|||||
Cp 83 AGGACTGCTAAGAACTTATGCTG 60

Search completed: Sun Mar 7 04:24:28 1999
Job time : 891 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mperch_un n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 04:28:50 1999; Maspar time 883.30 Seconds
1479.490 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-5
Description: (1-553) from US08962094.seq
Perfect Score: 553
N.A. Sequence: 1 GAATTGGGCTCGAGCGGCTC.....AAAAAARGGCGGCGCGC 553
Comp: CTTAAGCCGAGCTCGCGAG.....TTTTTTTCCCGCGCGCGC

Scoring table: TABLE default
Gap 6

Nmatch SMD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 508

Listing first 1000 summaries

Database:

emb157

1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_com 8:em_ov 9:em_ov 10:em_pat 11:em_ph 12:em_pl

Database:

genbank110

16:gb_ba1 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_com 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl1 25:gb_pl2 26:gb_pl
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_v1

Statistics: Mean 11.381; Variance 19.039; scale 0.598

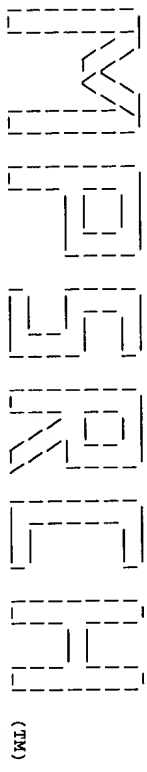
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
------------	-------	-------------	--------	----	-------------	-----------

No matches found.

Search completed: Sun Mar 7 04:48:00 1999
Job time : 1150 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_un n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:04:17 1999; MasPar time 86.58 Seconds
868.945 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-5
Description: (1-553) from US08962094.seq
Perfect Score: 553
N.A. Sequence: 1 GAATTCGGCTCGAGCGGC.....AAAAAAAAARGGCGCGCCGC 553
Comp: CTTAGCCGAGCTCGCCGAG.....TTTTTTTTTCCCGCGCGCG

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 508
Listing first 1000 summaries

Database: n-geneseg32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 8.751; Variance 8.509; scale 1.028

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query			
No.	Score	Match Length DB ID	Description Pred. No.

No matches found.

Search completed: Sun Mar 7 05:07:52 1999
Job time : 215 secs.


```
|||||
OY 219 GCTGTCGACACACGCGACGACCTGCTCTCTACCATCGACACACCGCTGCTTACC 278
DB 241 ACTGCTCGTAAAGACATTCACATTTTACCACCAATGGTTGGGAGATCTCCGATGGTGA 300
OY 279 ACTGCTCGTAAAGACATTCACATTTTACCACCAATGGTTGGGAGATCTCCGATGGTGA 338
DB 301 GGTGTCCCTGATGAGATGAGTGTGATGCTTTCGCAATNGTGACACATATTCATGC 360
OY 339 GGTGTCCCTGATGAGATGAGTGTGATGCTTTCGCAATNGTGACACATATTCATGC 397
DB 361 TTCCGTGATTCATCACTACTACCTTGCCTACGATATCCCTTATCTTAATCAG 420
OY 398 TTCCGTGATTCATCACTACTACCTTGCCTACGATATCCCTTATCTTAATCAG 457
DB 421 TTATTTCTTTCCAAA 436
OY 458 TTATTTCTTTCCAAA 473

RESULT 2
LOCUS W72838 432 bp mRNA EST 16-OCT-1996
DEFINITION zdf7912.s1 Soares fetal heart NBH19W Homo sapiens cDNA clone
344806 3' similar to contains element MER40 repetitive element ;
ACCESSION W72838
NID G1382814
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 432)
AUTHORS Hillier L., Clark N., Dubuque T., Ellston K., Hawkins M.,
Holman M., Hulman M., Kucaba T., Le M., Lennon G., Matra M.,
Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
Trevaskis E., Waterston R., Williamson A., Wohlmann P., and
Wilson R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 508 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 394.
location/Qualifiers
1..432
/organism="Homo sapiens"
/Note="Organ: heart; Vector: p773D (Pharmacia) with a
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTCACATCGAGTGGAGCGCCGCGATCTTTTCTTTTCTTTT 3')
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M. Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBH19W.
/db_xref="taxon:9606"
/clone="344806"
/clone.lib="Soares fetal heart NBH19W"
/sex="unknown"
/dev_stage="19 weeks"
```

```
|||||
BASE COUNT 128 a 78 c 124 g 102 t
ORIGIN
Query Match 76.98; Score 425; DB 21; Length 432;
Best Local Similarity 99.8%; Pred. No. 1,14e-221;
Matches 431; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

DB 1 GTTGCTCATAGTATTTTATTTTAAAGAAATTAATGATAGATAAAGGGATAT 60
CP 494 GTTGCTCATAGTATTTTATTTTAAAGAAATTAATGATAGATAAAGGGATAT 435
DB 61 CGTAGGCAAGGTAGTATGTTGATGAATACAGAGAACATGAATAGTGTGACCAATTG 120
CP 434 CGTAGGCAAGGTAGTATGTTGATGAATACAGAGAACATGAATAGTGTGACCAATTG 375
DB 121 CAGAGACTCAAGCTGATTCATCTCAGGAGACACCTCTACCATTCGGAGATCCCAAC 180
CP 374 CAGAGACTCAAGCTGATTCATCTCAGGAGACACCTCTACCATTCGGAGATCCCAAC 315
DB 181 CCATTTGGGTAAAGCTGGAATGCTCTTACGAGCAGTGTGTAGACAGCGGTGTCAGT 240
CP 314 CCATTTGGGTAAAGCTGGAATGCTCTTACGAGCAGTGTGTAGACAGCGGTGTCAGT 255
DB 241 GGTAGAGACAGCAGTGTGTCAGTGTGTCAGCAGCAGTGTGTCAGCAGCAGTGTG 300
CP 254 GGTAGAGACAGCAGTGTGTCAGTGTGTCAGCAGCAGTGTGTCAGCAGCAGTGTG 195
DB 301 ATCATCAGAGAGACAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTG 360
CP 194 ATCATCAGAGAGACAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTG 136
DB 361 GGGCAGAGACAGAAAGATGGAATCTCCAGAGTACAGAGCTGTAGAACTTCATGG 420
CP 136 GGGCAGAGACAGAAAGATGGAATCTCCAGAGTACAGAGCTGTAGAACTTCATGG 76
DB 421 TGTGACATGAA 432
CP 75 TGTGACATGAA 64

RESULT 3
LOCUS A1139456 402 bp mRNA EST 23-SEP-1998
DEFINITION qc20901.x1 Soares fetal heart NBH19W Homo sapiens cDNA clone
IMAGE:1710192 3' similar to contains element MER40 repetitive
element ; mRNA sequence.
ACCESSION A1139456
NID G3645428
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Euthera;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 402)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 398.
location/Qualifiers
1..402
/organism="Homo sapiens"
/Note="Organ: heart; Vector: p773D (Pharmacia) with a
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
```


AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fulderer, R.A., White, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., Milte, O., Sutcliffe, G., Blake, J.A., Brandon, R.C., Man-Wei, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodet, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Benharik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., DiRke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Korak, D.L., Kunsch, C., Hungjun, J., Li, H., Meisner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Rubin, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.

TITLE
JOURNAL
MEDLINE
COMMENT
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgl/hgl.html>)
Seq primer: M13 Reverse.

FEATURES

source
1. 294
Location/Qualifiers
/organism="Homo sapiens"
/note="Organ: skin; Vector: pbluescript SK-; Site_1: EcoRI; Site_2: XhoI"
/db_xref="ATCC (lnhost):141700"
/db_xref="taxon:9606"
/clone_idb="Fetal skin"
/tissue_type="epithelium"
/cell_type="epithelial cell"
/dev_stage="fetus"
1. >294
mRNA
BASE COUNT 60 a 91 c 58 g 82 t 3 others
ORIGIN

Query Match 52.3%; Score 289; DB 25; Length 294;
Best Local Similarity 98.6%; Pred. No. 6,15e-142;
Matches 290; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 1 CTTGAGACATTTTNTGTCCTGATCTTCTGAGTCCACACCATGAGTTTCTAGC 60
|||||
Oy 32 CTTTGAACATTTTGTCTGCTCCGATCTTATGTCACACCATGAGTTTCTAGC 91
|||||
Db 61 AGCTCTGATCTTGGAGATTTCATCTTNTGTCCTGCGCCAGAAATCCGACACAGC 120
|||||
Oy 92 AGCTCTGATCTTGGAGATTTCATCTTNTGTCCTGCGCCAGAAATCCGACACAGC 151
|||||
Db 121 TGTCTCAGCTGACAGCTATCCAGCTAGTGGTCTGATGATGAAAGCCCGATGCTGA 180
|||||
Oy 152 TGTCTCAGCTGACAGCTATCCAGCTAGTGGTCTGATGATGAAAGCCCGATGCTGA 211
|||||
Db 181 AACCACTCTCTGCAACCACTGCAACCACTGCTCTCTACCACTGCAACCACTGCTGC 240
|||||
Oy 212 AACCACTCTCTGCAACCACTGCAACCACTGCTCTCTACCACTGCAACCACTGCTGC 271
|||||
Db 241 TTCTACCACTCTCTGTAAGACATTCAGTTTACCAAAATGGGTTGGGATCT 294
|||||
Oy 272 TTCTACCACTCTCTGTAAGACATTCAGTTTACCAAAATGGGTTGGGATCT 325
|||||

RESULT

LOCUS 6 A1127172 443 bp mRNA EST 11-SEP-1998
DEFINITION g959a08.x1 Soares_fetal_heart_NBH19W Homo sapiens cDNA clone IMAGE:1708214 3', mRNA sequence.

ACCESSION
G1527172
KEYWORDS
93595686
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 443)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd, RT from Amersham.

FEATURES

source
1. 443
Location/Qualifiers
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'] TGTTCACCACTGATGAGGAGGCGGCGCATCTTTTCTTTTCTTTT [3'] double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBH19W."
/db_xref="taxon:9606"
/clone_image="IMAGE:1708214"
/clone_idb="Soares_fetal_heart_NBH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 129 a 80 c 129 g 103 t 2 others
ORIGIN

Query Match 52.3%; Score 289; DB 17; Length 443;
Best Local Similarity 94.8%; Pred. No. 6,15e-142;
Matches 364; Conservative 0; Mismatches 9; Indels 11; Gaps 11;

Db 60 TAAAGGAGATATGCGAGGCAAGTAAAGTGAATGAATACAGATTAACATGCA 119
|||||
Cp 446 TAAAGGAGATAT-CGTAGCGAAGTAAAGTAAAGTGAATGAATACAGATTAACATGCA 392
|||||
Db 120 ATAGTCTGTGACACAATTCGAAAGAGCTACAGCTGATTCATCATCAGGACACATCTC 179
|||||
Cp 391 ATAGT-CTGTGAC-CAATTCGAAAGAGCTCA-AGCTGATTCATC-TCAGGACACACA-CTC 337
|||||
Db 180 TACCTATTCGGGGAATCCCAACCATTTGGGTAAACTGGAATGCTCTTTACAGACAGTG 239
|||||
Cp 336 TACC-ATTCTGGGGAATCCCAACCATTTGGGTAAACTGGAATGCTCTTTACAGACAGTG 278
|||||
Db 240 GTAGAGAGAGGCTGTGTCAGTGTGAGAGACAGAGTGGCAGTGTGTCAGACGA 299
|||||
Cp 277 GTAGAGAGAGGCTGTGTCAGTGTGAGAGACAGAGTGGCAGTGTGTCAGACGA 218
|||||
Db 300 GTGTTTCAGCATAGGAGGCTTCATCATCAGACAGACAGTATGATGATGATGATGAT 359
|||||
Cp 217 GTGTTTCAGCATAGGAGGCTTCATCATCAGACAGACAGTATGATGATGATGATGAT 158
|||||
Db 360 GGAGCAGCTGTGTCGATTTGTGGGAGAGACCAAGAAAGATGGAACTGCCAGAGTACC 419
|||||

Mon Mar 8 15:42:06 1999

US-08-962-094-5.rst

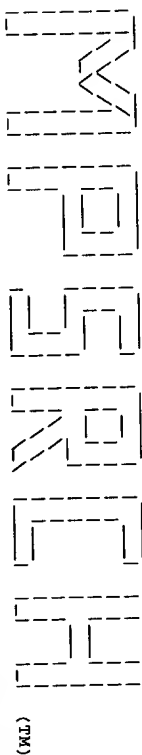
Page 5

CP 157 GGAGCAGCTGTGTGGATTCTGGGAGAGACCAAGAAAGATGGAACCTCCCAAGAGTACC 98

DB 420 AGGACTGCTAAGAGACTTAATGCTG 443

CP 97 AGGACTGCTAAGAGACTTCATGCTG 74

Search completed: Sun Mar 7 05:03:54 1999
Job time : 932 secs.



 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mserch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:12:24 1999; MasPar time 9.19 Seconds
 68.651 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-17
 Description: (1-39) from US08962094.pep
 Perfect Score: 213
 Sequence: 1 AONPTTAAPADTYPATGPDDEAPDAETAAATTATTTAA 39

Scoring table: PAM 150
 Gap 11

Searched: 131922 segs, 16180660 residues

Post-processing: Minimum Match 50%
 Listing first 1000 summaries

Database: a-geneseq32
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29

Statistics: Mean 19.949; Variance 100.669; scale 0.198

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.

No matches found.

Search completed: Sun Mar 7 05:14:27 1999
 Job time : 123 secs.

W P E R E H (TM)

 Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSPCH_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:11:00 1999; MasPar time 4.13 Seconds
 353.996 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-17
 Description: (1-39) from US08962094.pep
 Perfect Score: 213
 Sequence: 1 AQNPTTAAPADTYPATGPADEAPDAETTAATTATTAA 39
 Scoring table: PAM 150
 Gap 11

Searched: 116738 seqs, 37463448 residues

Post-processing: Minimum Match 50%
 Listing first 1000 summaries

Database: p1r58
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 28.767; Variance 75.764; scale 0.380
 Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.

No matches found.

Search completed: Sun Mar 7 05:12:02 1999
 Job time : 62 secs.

Mon Mar 8 15:41:44 1999

US-08-962-094-17.rsp

Page 1

MIRAGE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_gp protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:08:14 1999; Maspar time 3.87 Seconds
270.715 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-17
Description: (1-39) from US08962094.pep
Perfect Score: 213
Sequence: 1 AQNPTTAAPADTYPATGPADDEAPDAETTAATTATTAA 39

Scoring table: PAM 150
Gap 11

Searched: 74019 seqs, 26840295 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: swiss-prot35
1:swissprot

Statistics: Mean 29.901; Variance 73.335; scale 0.408

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.

No matches found.

Search completed: Sun Mar 7 05:09:15 1999
Job time : 61 secs.

WIPSON (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:09:35 1999; MasPar time 8.40 Seconds

231.126 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-17
Description: (1-39) from US08962094.pep
Perfect Score: 213
Sequence: 1 AQNPTTAAPADTYPATGPADDEAPDAETTAATTAATA 39

Scoring table: PAM 150
Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: sptrembl6
1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 28.755; Variance 72.941; scale 0.394

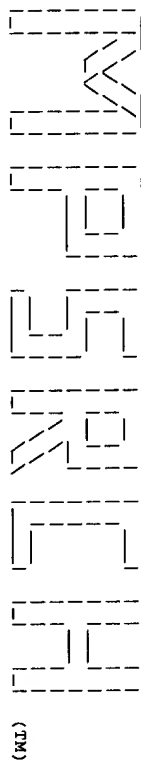
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
---------------	-------	----------------	--------	----	----	-------------	-----------

No matches found.

Search completed: Sun Mar 7 05:10:41 1999
Job time : 66 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mparch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:20:12 1999; Maspar time 4.17 Seconds
151.362 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-18
Description: (1-39) from US08962094.pep
Perfect Score: 252
Sequence: 1 TTATTAAPTTATTAATTARKDIPVLPKWGDLPNRGVC 39

Scoring table: PAM 150
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: a-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 23.465; Variance 92.489; scale 0.254

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	ID	Description	Pred. No.
No.	Score	Match	Length	DB

No matches found.

Search completed: Sun Mar 7 05:22:07 1999

Job time : 115 secs.

Mon Mar 8 15:41:47 1999

US-08-962-094-18.rpt

Page 1

WORLD (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPerch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:18:28 1999; Maspar time 5.95 Seconds
245.440 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-18
Description: (1-39) from US08962094.pep
Perfect Score: 252
Sequence: 1 TTATTAAPTTTATTAATASTARKDIPVLPKWGDLPNGRVC 39

Scoring table: PAM 150
Gap 11

Searched: 116738 seqs, 37463448 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: p1r58
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 32.408; Variance 66.743; scale 0.486

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB	ID	Description	Pred. No.
---------------	----------------	--------	----	----	-------------	-----------

No matches found.

Search completed: Sun Mar 7 05:19:49 1999
Job time : 81 secs.

 WIDEORLH (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch_p protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:14:47 1999; MasPar time 2.82 Seconds
 370.807 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-18
 Description: (1-39) from US08962094.pep
 Perfect Score: 252
 Sequence: 1 TTATTAAPTTATTASTARKDIPVLPKWGDLPNGRVC 39

Scoring table: PAM 150
 Gap 11

Searched: 74019 segs, 26840295 residues

Post-processing: Minimum Match 50%
 Listing first 1000 summaries

Database: swiss-prot35
 1:swissprot

Statistics: Mean 33.471; Variance 60.737; scale 0.551

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.

No matches found.

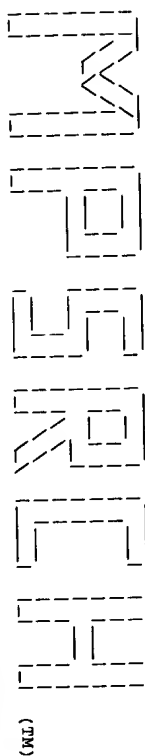
Search completed: Sun Mar 7 05:16:17 1999
 Job time : 90 secs.

10

Mon Mar 8 15:41:48 1999

US-08-962-094-18.rsp

Page 1



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mp501H protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:16:40 1999; MasPar time 9.93 Seconds
195.650 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-18
Description: (1-39) from US08962094.pep
Perfect Score: 252
Sequence: 1 TTATTAAPTTATTAASTARKDIPVLPKWGDLPNGRYC 39

Scoring table: PAM 150
Gap 11

Searched: 165420 segs, 49795644 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: sptrembl6
1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

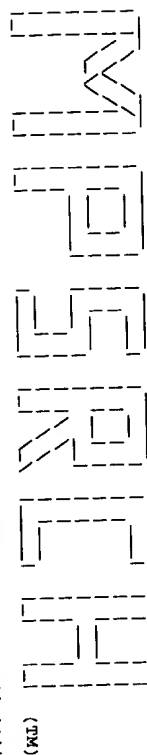
Statistics: Mean 32.963; Variance 61.313; scale 0.538
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred. No.
------------	-------------	-------	--------	----	-------------	-----------

No matches found.

Search completed: Sun Mar 7 05:18:06 1999
Job time : 86 secs.



 Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_LP protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:26:24 1999; Maspar time 8.97 Seconds
 37.882 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-19
 Description: (1-21) from US08962094.pep
 Perfect Score: 166
 Sequence: 1 ARKDIPVLPRKWGDLNCRVC 21

Scoring table: PAM 150
 Gap 15

Searched: 131922 segs, 16180660 residues

Post-processing: Minimum Match 50%
 Listing first 1000 summaries

Database: a-geneseq32
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29

Statistics: Mean 21.655; Variance 69.163; scale 0.313

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.

No matches found.

Search completed: Sun Mar 7 05:28:03 1999

Job time : 99 secs.

WATERMAN (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msearch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:25:03 1999; Maspar time 4.79 Seconds
164.193 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-19
Description: (1-21) from US08962094.pep
Perfect Score: 166
Sequence: 1 ARKDIPVLPKMWGDLPGNRVC 21

Scoring table: PAM 150
Gap 15

Searched: 116738 segs, 37463448 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: p1r58
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 29.205; Variance 46.987; scale 0.622

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred. No.
------------	-------------	-------	--------	----	-------------	-----------

No matches found.

Search completed: Sun Mar 7 05:26:01 1999
Job time : 58 secs.

MIPSREL
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

Mpsrch_p protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:22:30 1999; Maspar time 3.54 Seconds
159.000 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-19
Description: (1-21) from US08962094.pep
Perfect Score: 166
Sequence: 1 ARKDIPVLPMWGDLPNGRVC 21

Scoring table: PAM 150
Gap 15

Searched: 74019 seqs, 26840295 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: swiss-prot35
1:swissprot

Statistics: Mean 30.068; Variance 42.251; scale 0.712

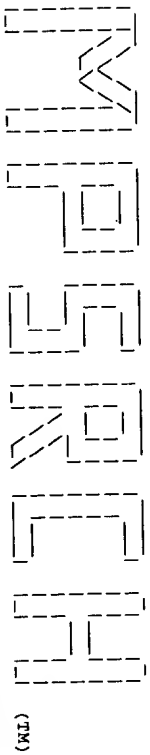
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.
------------	-------------	--------	----	-------------	-----------

No matches found.

Search completed: Sun Mar 7 05:23:23 1999
job time : 53 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mparch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:23:45 1999; Maspar time 8.49 Seconds
123.224 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-19
Description: (1-21) from US08962094.psp
Perfect Score: 166
Sequence: 1 ARKDIPVLPKRWGDLPGNGRC 21

Scoring table: PAM 150
Gap 15

Searched: 165420 segs, 49795644 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: sptrembl6
1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_orfanelle
9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

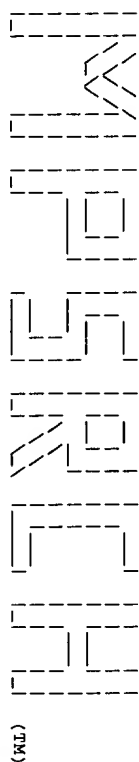
Statistics: Mean 29.368; Variance 42.447; scale 0.692
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred. No.

No matches found.

Search completed: Sun Mar 7 05:24:42 1999
Job time : 57 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sun Mar 7 05:32:48 1999; MasPar time 8.81 Seconds
38.578 Million cell updates/sec

Tabular output not generated.

Title: >US-08-962-094-20
Description: (1-21) from US08962094.pep
Perfect Score: 125
Sequence: 1 AAPADTYPATGPADDEAPDAE 21

Scoring table: PAM 150
Gap 15

Searched: 131922 segs, 16180660 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: a-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 18.207; Variance 65.743; scale 0.277

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred. No.
No matches found.						

No matches found.

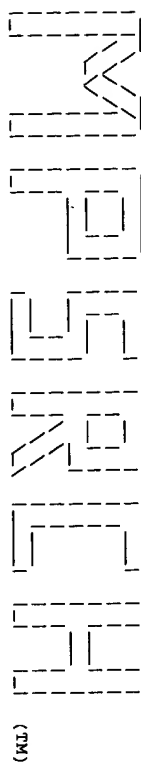
Search completed: Sun Mar 7 05:34:39 1999

Job time : 111 secs.

3

3

11



Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mserch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Sun Mar 7 05:28:24 1999; Mspartime 3.92 Seconds
143.779 Million cell updates/sec
Tabular output not generated.

Title: >US-08-962-094-20
Description: (1-21) from US08962094.pep
Perfect Score: 125
Sequence: 1 AAPADTYPATGPADDEAPDAE 21

Scoring table: PAM 150
Gap 15

Searched: 74019 segs, 26840295 residues

Post-processing: Minimum Match 50%
Listing first 1000 summaries

Database: swiss-prot35
1:swissprot

Statistics: Mean 26.391; Variance 38.782; scale 0.681

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	66	52.8	690	1 RHO_MICLU	TRANSCRIPTION TERMINAT	1.21e-01
2	63	50.4	825	1 5E5_RAT	5E5 ANTIGEN.	4.36e-01

Note: Post-processor removed 998 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1
ID RHO_MICLU STANDARD; PRT; 690 AA.
AC P52154;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE TRANSCRIPTION TERMINATION FACTOR RHO.
GN RHO.
OS MICROCOCCUS LUTEUS (MICROCOCCUS LYODEIKTICUS).
OC PROKARYOTA; FIRMICUTES; COCCI; MICROCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5 AND 289-297.
RC STRAIN-EM;
RX MEDLINE: 96132802.
RA NOMATZKE W.L., RICHARDSON J.P.;

RL J. BIOL. CHEM. 271:742-747(1996).
RN [2]
RP SEQUENCE OF 205-690 FROM N.A.
RC STRAIN-EM;
RX MEDLINE: 94327472.
RA OPPERMAN T., RICHARDSON J.P.;
RL J. BACTERIOL. 176:5033-5043(1994).
RN [3]
RP REVISION TO 500.
RC STRAIN-EM;
RA NOMATZKE W.L.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM
CC THAT INVOLVES RHO BINDING TO THE NASCENT RNA, ACTIVATION OF RHO'S
CC RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE RNA FROM THE
CC DNA TEMPLATE. RNA-DEPENDENT NTPASE WHICH UTILIZES ALL FOUR
CC RIBONUCLEOSIDE TRIPHOSPHATES AS WELL AS DATP AS SUBSTRATES, BUT
CC HAS A SIGNIFICANT LOWER ACTIVITY WITH CTP.
CC -1- SUBUNIT: HOMOHETEROMER (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNP).
CC EMBL: L27277; G166540;
DR TRANSCRIPTION TERMINATION; HELICASE; ATP-BINDING; RNA-BINDING.
KW INT_MET 0
FT DOMAIN 30 35 RNA-BINDING (RNP2) (BY SIMILARITY).
FT DOMAIN 318 321 RNA-BINDING (RNP1) (BY SIMILARITY).
FT NP_BIND 441 448 ATP (POTENTIAL).
FT CONFLICT 291 291 G -> P (IN AA SEQUENCE).
SQ SEQUENCE 690 AA; 75030 MM; 93AB0634 CRC32;

Query Match 52.8%; Score 66; DB 1; Length 690;
Best Local Similarity 47.6%; Pred. No. 1.21e-01;
Matches 10; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

DB 71 AAPAETAPAAASSEDAPAAE 91
QY 1 AAPADTYPATGPADDEAPDAE 21

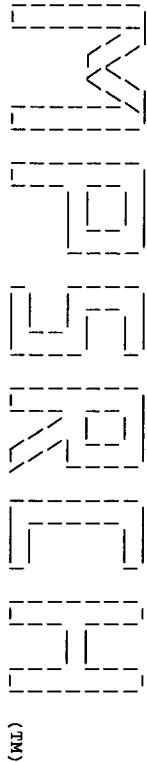
RESULT 2
ID 5E5_RAT STANDARD; PRT; 825 AA.
AC Q63003;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE 5E5 ANTIGEN.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WISTAR; TISSUE-BRAIN;
RX MEDLINE: 96015159.
RA SUZUKI E., KOJIMA N., YOSHIMURA K., UYEMURA K., OBATA K., AKAGAWA K.;
RL J. BIOCHEM. 118:122-128(1995).
CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
DR EMBL: D37934; G531261;
KW DNA-BINDING; NUCLEAR PROTEIN; ANTIGEN.
SQ SEQUENCE 825 AA; 86831 MM; BC01DIEF CRC32;

Query Match 50.4%; Score 63; DB 1; Length 825;
Best Local Similarity 37.5%; Pred. No. 4.36e-01;
Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

DB 138 SPEGAPGGGPAEEG 153
QY 2 AAPADTYPATGPADDEA 17

Search completed: Sun Mar 7 05:29:13 1999
Job time : 49 secs.

10



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msearch_p protein - protein database search, using Smith-Waterman algorithm
Run on: Sun Mar 7 05:29:33 1999; Maspac time 5.15 Seconds
Tabular output not generated.
Title: >US-08-962-094-20
Description: (1-21) from US08962094.1.ppt
Perfect Score: 125
Sequence: 1 AAPADTYPATGPADDEAPDAE 21
Scoring table: PAM 150
Gap 15
Searched: 165420 segs, 49795644 residues
Post-processing: Minimum Match 50%
Listing first 1000 summaries
Database: sprenb16
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus
Statistics: Mean 25.336; Variance 37.686; scale 0.672
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
SUMMARIES
Result No. Score Match Length DB ID Description Pred. No.
1 64 51.2 1463 14 069298 MAJOR IMMEDIATE EARLY 3.62e-01
2 63 50.4 801 5 023635 SIMILAR TO LONG TANDEM 5.56e-01
3 63 50.4 3158 2 052819 PCZA363.3. 5.56e-01
Note: Post-processor removed 997 summaries from list due to search parameters chosen.
ALIGNMENTS
RESULT 1
ID 069298 PRELIMINARY; PRT; 1463 AA.
AC 069298;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE MAJOR IMMEDIATE EARLY PROTEIN.
GN
OS
OC GALVAD HERPESVIRUS TYPE 1.
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; GAMMAHERPESVIRIDAE.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-SA-2;
RX MEDLINE: 95282504.
RA JOHNSON M.A., TYACK S.G., PRIDEAUX C., KONGSUTAN K., SHEPPARD M.;
RL VIRUS RES. 35:193-204(1995).
DR EMBL; L32139; G493598;
SQ SEQUENCE 1463 AA; 158701 MW; CA15F280 CRC32;
Query Match 51.2%; Score 64; DB 14; Length 1463;
Best Local Similarity 64.3%; Pred. No. 3.62e-01;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 947 SGPAGSYATGPAD 960
QY 1 AAPADTYPATGPAD 14
RESULT 2
ID 023635 PRELIMINARY; PRT; 801 AA.
AC 023635;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE SIMILAR TO LONG TANDEM REPEAT REGION OF STALIDASE.
GN ZK84.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACCELLOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., BONFIELD J.,
RA BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A., CRAXTON M.,
RA DEAR S., DU Z., DORBIN R., FAVELLO A., FULTON L., GARDNER A., GREEN P.,
RA HAWKINS T., HILLIER L., JIER M., JOHNSTON L., JONES M., KERSHAW J.,
RA KIRSTEN J., LAISTER N., LATREILLE P., LIGHTNING J., LLOYD C.,
RA MCMURRAY A., MORTIMER B., O'CALLAGHAN M., PARSONS J., PERCY C.,
RA RIFKEN L., ROOFA A., SANDERS D., SHOWNKEEN R., SMALDON N., SMITH A.,
RA SONNHAMMER E., STADEN R., SULLSTON J., THIERRY-MIEG J., THOMAS K.,
RA VAUDIN M., VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA WILKINSON-SPROAT J., WOHLDMAN P.;
RL NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA KIRSTEN J.;
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U23181; G726436;
SQ SEQUENCE 801 AA; 77123 MW; 7288A89D CRC32;
Query Match 50.4%; Score 63; DB 5; Length 801;
Best Local Similarity 42.9%; Pred. No. 5.56e-01;
Matches 9; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
Db 411 SAAAEETPAPAPAEETPAPE 431
QY 1 AAPADTYPATGPADDEAPDAE 21
RESULT 3
ID 052819 PRELIMINARY; PRT; 3158 AA.
AC 052819;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE PCZA363.3.
OS AMYCOLATOPHYS ORIENTALIS.
OC EUBACTERIA; FIRMICUTES; ACTINOMYCETES; PSEUDONOCARDIACEAE;
OC AMYCOLATOPHYS.
RN [1]
RP SEQUENCE FROM N.A.
RA VAN WAGENINGEN A., KIRKPATRICK P., WILLIAMS D., HARRIS B., KERSHAW J.,
RA LENNARD N., JONES M., JONES S., SOLENERG P.;

RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL AJ223999; E1231243; -

DR PROSITE: PS00455; AMP_BINDING; 3.

SQ SEQUENCE 3158 AA; 336315 MW; E162473A CRC32;

Query Match 50.4%; Score 63; DB 2; Length 3158;

Best Local Similarity 47.1%; Pred. NO. 5.56e-01;

Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Db 2430 ENYPTPAAGDEAADAQ 2446

OY 5 DTYPATCPADDEAPDAE 21

Search completed: Sun Mar 7 05:30:51 1999
Job time : 78 secs.